# Case Study: Analyze Video Game Sales

## Data File

Dataset contains list of video games that sold more than 100,000 copies.

Dataset variables:

Rank - Ranking of overall sales

Name - The games name

Platform - Platform of the games release (i.e. PC,PS4, etc.)

Year - Year of the game's release

Genre - Genre of the game

Publisher - Publisher of the game

NA\_Sales - Sales in North America (in millions)

EU\_Sales - Sales in Europe (in millions)

JP\_Sales - Sales in Japan (in millions)

Other\_Sales - Sales in the rest of the world (in millions)

Global\_Sales - Total worldwide sales.

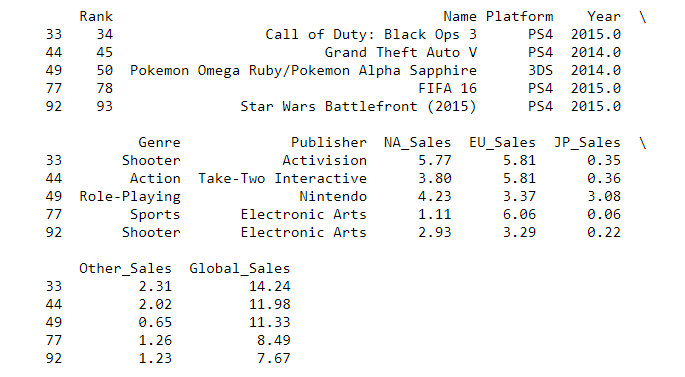
Data location: <https://www.kaggle.com/gregorut/videogamesales>

## Overview

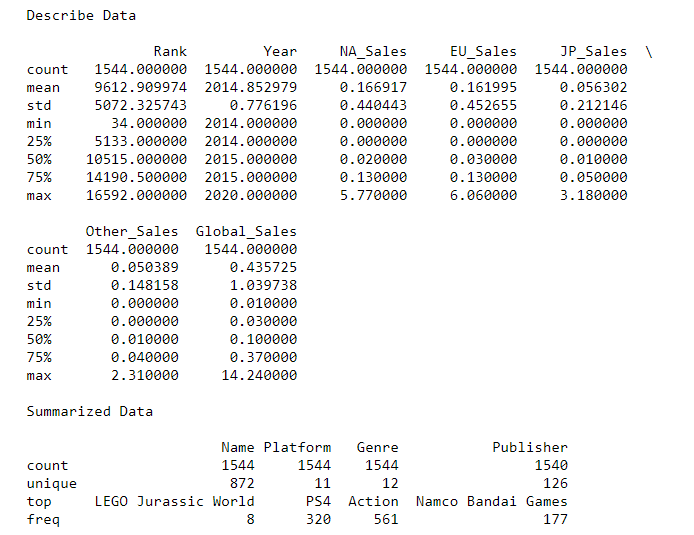
This analysis will look at video game sales data from various genres and platforms to see what can be determined for global sales. This analysis will focus on newer data to determine where the industry is going. Likewise, some platforms will be eliminated based on the fact that they are being phased out.

## Step-by-Step Instructions

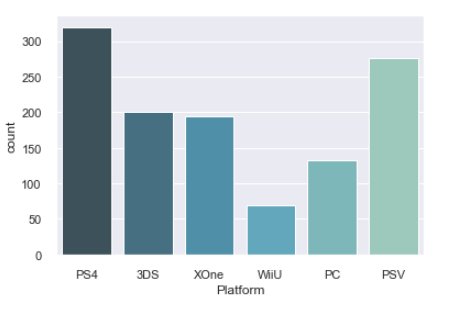
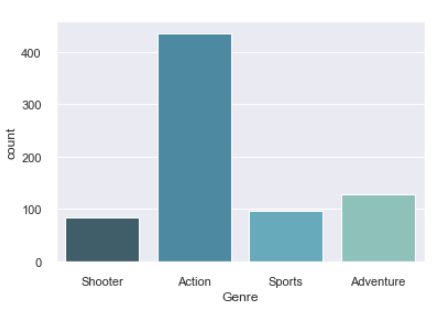
1. Load the data from “vgsales.csv” file into a DataFrame.
2. Display the dimensions of the file.  
     
   *The dimension of the data is: (16598, 11)*
3. Clean up the data. Since this analysis is intended to help predict a future successful game we only want to look at data from the last few years (2014 – present).
4. Display the dimensions of files.  
     
   *The dimension of the new data is: (1544, 11)*
5. Display the first few rows of the new dataset.



1. Display the summary of the data.

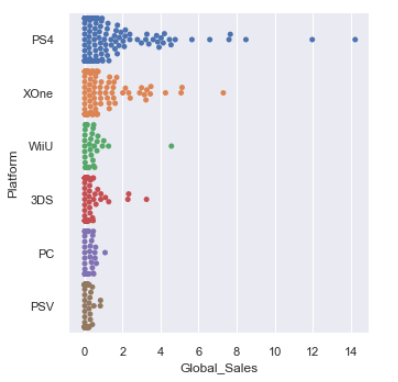
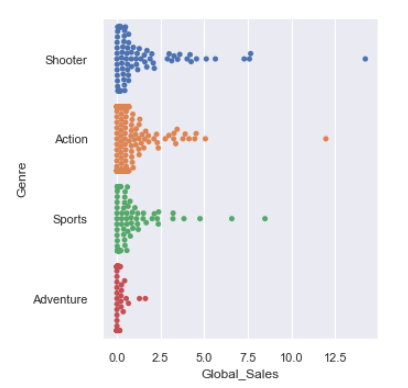


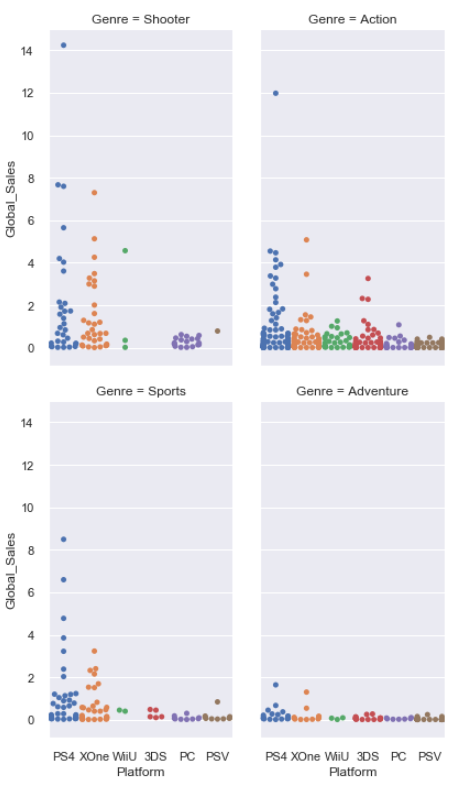
1. Plot histogram for genre and platform.

* 1. Platform histogram conclusion – After reviewing the histogram for the platforms, the PSP, DS, Wii, X360, and PS3 were eliminated and only the above ones remain. This was due to lower sales or phased out systems.
  2. Genre histogram conclusion – Various genres were eliminated due to lower sales performance; only the higher ranked genres were kept and displayed above.

1. Plot Sales vs. Genre/Platform.



Of the remaining genres and platforms we can conclude that action, shooter, and sports style games are the best in terms of sales.

## Future Considerations

After performing the above analysis I determined that the dataset selected was insufficient. An additional dataset that incorporates more features like user score and critics score should have been used. The above dataset provides data on sales that have occurred, while this can as an indication of future sales but there could be instance where people buy a game and hate… leading not buying in the future. Additional analysis should be applied to demonstrate trends in year vs sales to help assist in future.

Additional dataset: <https://www.kaggle.com/ashaheedq/video-games-sales-2019>